Convergence in West German Regional Unemployment Rates

Christian Bayer*
Falko Juessen†‡
University of Dortmund

This version: February 15, 2006

Abstract

Differences in regional unemployment rates are often used to describe regional economic inequality. This paper asks whether changes in regional unemployment differences in West Germany are persistent over time. Understanding the persistency of regional unemployment differences helps us to asses how effective regional policy can be. While univariate tests suggest that changes in unemployment differences are persistent, more powerful panel tests lend some support to the hypothesis that regional unemployment rates converge. However, these tests reveal a moderate speed of convergence at best. Since there is a structural break following the second oil crisis, we also employ tests that allow for such a break. This provides evidence for both, convergence and quick adjustment to an equilibrium distribution of regional unemployment rates that is subject to a structural break.

JEL-classification: R23, J60

Keywords: stochastic convergence, unemployment, structural break, unit root

1 Introduction

The extensive literature on economic convergence between countries and regions focuses mostly on per capita income or other related income and productivity measures. This focus may be fruitfully extended to other areas in economics, as Quah (1996, p. 1354) has pointed out:

*University of Dortmund, Department of Economics, LS VWL (Microeconomics), 44221 Dortmund, Germany, Tel.: +49-231-755-3246 , email: Christian.Bayer@wiso.uni-dortmund.de

†University of Dortmund, Department of Economics, LS VWL (Macroeconomics), 44221 Dortmund, Germany, Tel.: +49-231-755-3291 , email: F.Juessen@wiso.uni-dortmund.de

‡We would like to thank two anonymous referees, Heinz Holländer, Carlos Usabiaga, Michael Roos, Christoph Alsleben, and Christiane Schuppert for their helpful comments. This research is supported by DFG under Sonderforschungsbereich 475.

> 'Certainly, understanding economic growth is important. But growth is only one of many different areas in economics where analyzing convergence sheds useful insight.'

Following Quah's general suggestion, this paper borrows techniques from the literature on growth convergence. We use these techniques to examine the evolution of regional disparities in unemployment rates within a country, a topic that has gained much attention since the seminal paper of Blanchard and Katz (1992).

Unemployment disparities are often perceived as persistent. They are at the heart of the 'regional problem' and in the focus of regional economic policy (Armstrong and Taylor, 2000). Thus, their persistence has attracted much attention.[1]

Persistency itself may reflect stable equilibrium differentials of regional unemployment rates or may be attributed to the fact that shocks to regional unemployment rates have long-lasting effects, see Martin (1997). Discriminating between these two cases is important because policy interventions are more likely to be effective in the latter case. On the contrary, if the differences reflect an equilibrium that has been stable over time, (short-term) policy interventions are less likely to change this stable equilibrium.

It is thus interesting in particular in the German context to study whether regional unemployment rates converge to the national average over time. As the federal government in Germany aims to reduce the gap between unemployment rates in East and West Germany by granting subsidies and by spending on public infrastructure, it is of importance to understand how fast convergence happens.

In order to understand how quickly unemployment rates converge, we employ aggregated annual data from the 'Mikrozensus' database on unemployment rates for the West German federal states during the period 1960-2002. We analyze the convergence issue using the stochastic approach that was proposed by Bernard and Durlauf (1995, 1996) and others. This means that our study characterizes the evolution of the gap between the unemployment rate in a specific federal state and the unemployment rate in Germany as a whole.

For the US, Blanchard and Katz (1992) have analyzed the dynamics of regional employment and unemployment. While they do not explicitly find evidence for stationarity of regional unemployment rates, they attribute this to a power problem of the tests they apply. Indeed, Decressin and Fatas (1995) and Obstfeld and Peri (1998) provide some evidence that regional unemployment disparities are a more persistent phenomenon in Europe than in the US. However, these results have recently been questioned by Rowthorn and Glyn (2003) who do find substantial persistence also in US regional unemployment

[1] See for example Decressin and Fatas (1995) or Obstfeld and Peri (1998).

rates. For the UK by contrast, Martin (1997) finds that regional unemployment shocks are only short-lived. Yet, he also finds that regional unemployment rates differ in the long run, which reflects a stable equilibrium distribution around the national average.

For Germany, our study is the first one analyzing the convergence of unemployment rates at the federal state level. There are a number of studies which examine the related issue of hysteresis for West German unemployment rates: Balz (1999), Belke (1996), Belke and Göcke (1996), Camarero and Tamarit (2004), Hansen (1991), and Reutter (2000). However, these studies analyze the absolute level of aggregate or regional unemployment rates and not relative unemployment rates as we do. As a consequence, these papers cannot shed much light on convergence.

The main results of our study are the following. While univariate techniques which do not account for structural breaks do not provide evidence for stochastic convergence in relative unemployment rates, more powerful panel-based methods allow us to infer that there is convergence. At the same time, the panel-based methods also suggest that the speed of convergence is slow. The estimated half-life of a shock to regional unemployment is at least 5.6 years.

However, this degree of persistence may be over-estimated. There is a structural break in the data following the second oil crisis as a graphical analysis reveals. In order to find out how strongly this break drives our previous results of non- or slow convergence, we subsequently apply an empirical framework that is robust to the existence of a structural break. This structural break is specified as an endogenously determined single level shift in the mean of the unemployment rate of each federal state relative to Germany as a whole. Under this specification, we can reject the null hypothesis that shocks to unemployment differences persist. Rather, the tests give evidence for conditional convergence in most regions. This conditional convergence means that regional unemployment rates converge up to a constant difference to the national average, but this difference is subject to a one-time permanent shift, which occurred following the second oil crisis. Moreover, allowing for a structural break, the estimated speed of convergence increases substantially, so that the estimated half-life goes down from 5.6 to less than 2 years on average. Consequently, persistency in regional unemployment disparities reflects an equilibrium to which the German economy adjusts quickly.

The remainder of this paper is organized as follows: Section 2 introduces the theoretical concepts. After describing the data in Section 3, we begin with a graphical analysis, which serves as a guideline for the rest of the paper. Section 4 provides the analysis of convergence on the basis of univariate and panel unit-root tests which do not account for structural breaks. This analysis is extended to the possibility of a structural break in Section 5.

Finally, Section 6 discusses our results and Section 7 concludes.

2 Theoretical concepts

When labor markets adjust towards equilibrium in the long run, there will be convergence of regional unemployment rates, because unemployed workers take jobs in other areas or because capital flows into low-wage regions to take advantage of lower labor costs (for details, see Blanchard and Katz, 1992). However, if the speed of adjustment is slow, unemployment disparities may arise during adjustment as a result of negative demand shocks affecting some regions more than others (Armstrong and Taylor, 2000).

We can test this theory of long-term convergence empirically by using Bernard and Durlauf's (1995, 1996) time-series approach. This approach focuses on the permanence of shocks to *relative* variables and uses a stochastic definition of convergence (Carlino and Mills, 1993).

The idea of Bernard and Durlauf's test for stochastic convergence can be explained as follows. Let ur_{it} and ur_{jt} be the unemployment rates of regions i and j at time t, respectively. Suppose that region i has a larger unemployment rate than region j initially, $ur_{i0} > ur_{j0}$. The gap in unemployment between the two regions is $ur_{it} - ur_{jt}$. Define I_t as the information set available at period t. Then, Definition 2 in Bernard and Durlauf (1996, p. 165) understands convergence as the equality of long-term forecasts at any fixed time. This means

$$\forall t : \lim_{s\to\infty} E(ur_{i,t+s} - ur_{j,t+s}|I_t) = 0. \tag{1}$$

Stochastic convergence implies that regional unemployment differences will always be transitory in the sense that the long-term forecast of the difference between any pair of regions tends to zero as the forecast horizon grows.

The important testable implication of this stochastic approach to long-term convergence is that convergence is present only if shocks to the unemployment differential are temporary. Hence, the disparities between regions should follow a stationary process, which means that ur_i and ur_j are cointegrated. Without stationarity, shocks to the relative variable lead to permanent differences.

As such, time-series tests of convergence have typically been implemented by using unit-root tests. For example, Carlino and Mills (1993) and Evans and Karras (1996) apply Dickey-Fuller type tests for the presence of a unit root in the relative variable. If the series has a unit root, shocks are permanent and there will be no convergence. Besides precluding stochastic trends (i.e. unit-roots), long-term convergence also precludes any deterministic trends in cross-regional differences. In fact, also the mean of the series of unemployment differences should be zero under the assumption of absolute convergence.

However, the hypothesis of stationarity *and* zero means might be too strict. As an example, we can consider regional amenities that lead to wage differentials which compensate workers for differences in the quality of life or for different regional price levels. If we assume additionally that there is a national unemployment insurance that pays a fixed unemployment benefit which is equal among regions, then there will be persistent differences in unemployment rates. Because wages are lower in regions rich of amenities, the equal unemployment benefit leads to higher rates of voluntary unemployment in the amenity-rich regions. Or to put it differently, the voluntarily unemployed would move to the amenity-rich regions in this simplistic setting. This results in stable differences between regional unemployment rates, whereas these differences just reflect disparities in economic fundamentals, such as differences in natural endowments.

In such a setting, regional economic policy that wants to reduce inequality would need to aim at shifting the equilibrium. However, it is unlikely that short-term policy interventions are actually effective for this purpose, if the equilibrium has been stable over the past.[2]

To capture this notion of stable long-term equilibrium differences, we define *conditional* convergence as

$$\forall t : \lim_{s \to \infty} E(ur_{i,t+s} - ur_{j,t+s} | I_t) = constant. \tag{2}$$

This means that ur_i and ur_j converge towards a (time-invariant) equilibrium differential. An empirical test for stochastic conditional convergence is again related to the time-series properties of relative unemployment rates. Conditional convergence implies that the series is (weakly) level-stationary but it is not required that the series has a zero mean.[3]

3 Data and graphical analysis

3.1 Data

We use data that is aggregated from the German 'Mikrozensus' database by the Federal Statistical Office. The Mikrozensus is an annual collection of household data for a representative sample of German households. The aggregated data is available from 1957 onwards to the scientific user, while the microdata is available only since 1989. For this reason, we use aggregated data at the federal state level, which is available since 1957.

Since there was virtually no unemployment in Germany during the late 1950s, we

[2] See Marston (1985) for a more elaborated theoretical underpinning of the equilibrium and disequilibrium perspective of regional unemployment disparities.

[3] We can consider the series generated by the autoregressive model $u_t = \phi + \rho u_{t-1} + \varepsilon_t$ as an example. This series is stationary if $|\rho| < 1$ and the intercept ϕ controls the mean of u_t through the relationship $E(u_t) = \mu = \phi/(1-\rho)$. If u is relative unemployment, we find conditional convergence if $\rho < 1$ and unconditional convergence if additionally $\phi = 0$.

restrict the data to the time-period 1960-2002. Moreover, West Berlin is excluded from the analysis because of its special status before German reunification.

The data contains information on the number of employed and on the number of unemployed persons for each federal state. In the Mikrozensus data, the term 'unemployed' refers to all people without employment contract who search for a job irrespective of being registered as unemployed or not at the German Federal Employment Agency. Therefore, the definition of unemployment in our data differs somewhat from the statistics of the German Federal Employment Agency, but is more similar to the definition of the unemployment rate used in other countries, in particular the US.[4]

Another central advantage of our data is that it spans a long period of time. The long period of time is important for our analysis for two reasons. Firstly, we want to find out whether relative unemployment rates exhibit some form of path dependency or converge alternatively. Obviously, observing the data over a long time span is crucial for such kind of analysis. Secondly, and even more importantly, the long time span allows us to assess whether regional unemployment disparities are subject to structural break over time. As it will turn out, allowing for structural breaks is important both, for our test results and even more so for the interpretation of the latter.

The unemployment rate (in percentage points) is defined as the number of unemployed divided by the labor force ('Erwerbspersonen') multiplied by 100. Labor force data was also derived from the Mikrozensus. According to the Mikrozensus definition, the labor force is the sum of the employed and the unemployed ('Erwerbstaetige' and 'Erwerbslose').

We denote the unemployment rate for federal state i by ur_i and the unemployment rate for Germany as a whole (without West-Berlin) by ur_{Ger}. Time indices are suppressed for notational convenience. For the period after German reunification, 1991-2002, the unemployment rate for Germany, ur_{Ger}, is calculated on the basis of the data from West German federal states only.

As explained in the previous section, stochastic convergence requires that relative unemployment rates follow a stationary process. We compute the relative unemployment rate u_i for federal state i as

$$u_i = ur_i - ur_{Ger} \tag{3}$$

The unemployment rate for Western Germany, ur_{Ger}, is selected as a reference. This reflects that unemployment rates for the different federal states do not evolve differently from the national average if they converge.[5]

[4] Annual data on *registered* unemployment at the federal state level is available only since 1974 (depending on the federal state).

[5] Using differences in logs or ratios of unemployment rates has the disadvantage that minor differences in unemployment rates and rounding errors get inflated by the low aggregate unemployment rates during

The typical testing strategy for convergence applies some linear model for u_{it} and a test for the presence of a unit root. Since unemployment rates are relative numbers and bounded between 0 and 100 percent, also relative unemployment rates are bounded between -100 and +100 percent. Hence, one may argue that taking literally the linear model for the differences implies that non-stationarity cannot take the form of a unit-root property of u_i. If u_i is non-stationary, this must stem from a more complicated non-linear dynamics that is path-dependent (see Amable et al., 1994 and 2005). For example, non-stationarity could originate from a threshold cointegrated process that is mean reverting outside a certain range and has a unit root inside this range. Whether one views such a process as stationary or non-stationary depends on the relevance of the reflecting boundaries. If the boundaries are close and hit often, describing the process as stationary is a good approximation. If by contrast the boundaries are hit seldom within the sample, we may best describe the sample as having a unit root, since the outside range looses relevance. Applying a unit-root test to such process reveals the importance of non-stationarity as a property to describe the sample. In other words, we understand the unit-root property as a sample property and the relevant question becomes how persistent is the process (Blanchard and Summers, 1986).

Keeping this in mind, we apply a linear framework and approximate a test for non-stationarity by means of a test for a unit root. We may attempt to assess the appropriateness of the linear framework beforehand by inspecting how close unemployment differences get to their boundary values. In fact, we find relative unemployment rates never to be close to the bounds -100 and +100 percent.

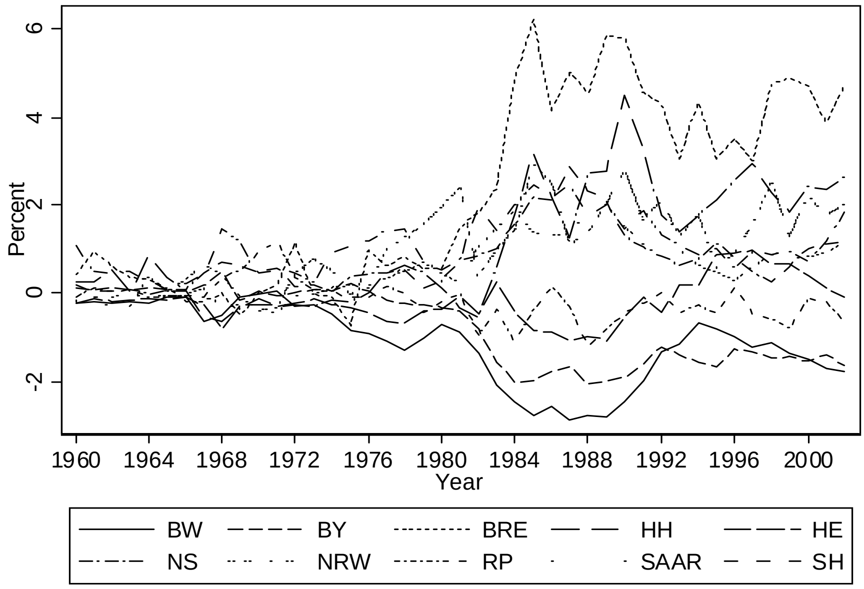
3.2 Graphical analysis

To get a first impression of the time-series characteristics of u_i, we display the series graphically. Figure 1 plots relative unemployment rates during the period 1960-2002.

It can be seen that the dispersion of unemployment rates has sharply increased in times of recessions (1966/67 and at the beginning of the 1980s) parallel to the increase in the aggregate unemployment rate. In the beginning of the 1960s unemployment was not a problem in Germany, in fact there was rather a shortage of labor, similarly there is not much of a difference in unemployment rates across states. After 1980 the situation is dramatically different, the dispersion of unemployment rates sharply increases with the general rise in unemployment rates. Thereafter, economic differences between the northern and southern part of Germany become apparent. Since the beginning of the 1980s, the North-German city-states Bremen and Hamburg have the highest relative unemployment rates, while Bayern and Baden-Württemberg have unemployment rates around 2 percentage points

the 1960s.

Figure 1: Relative unemployment rates in West Germany, 1960-2002



Data source: Mikrozensus, West Berlin excluded

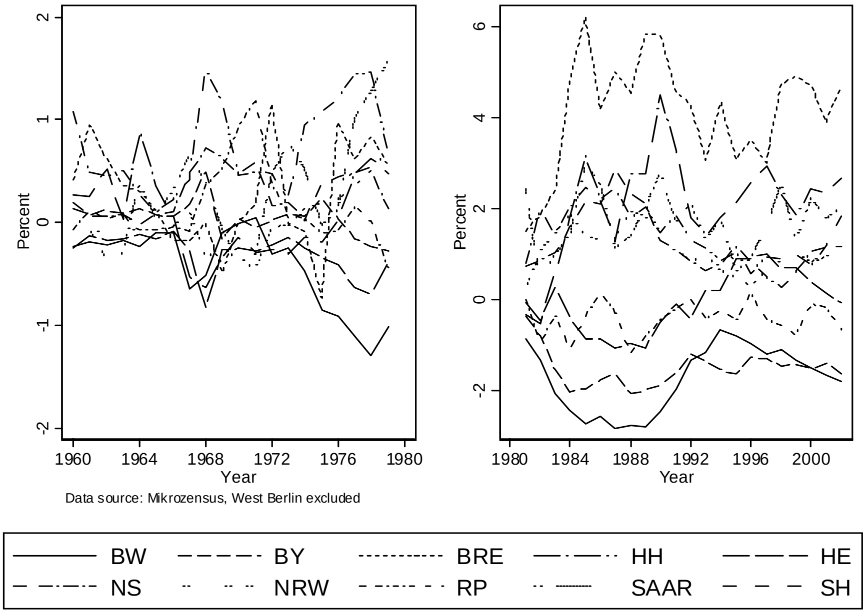
below the national average.

At first glance, this makes most of the series look non-stationary. However, splitting the sample in the period before and after 1980 shows that the lack of stationarity might just be due to a single structural break that occurs in the early 1980s after the second oil crisis. In order to illustrate this, Figure 2 displays the data for both sub-periods; one ranging from 1960-1979 and the second from 1980-2002 (Figure 2). The series look more stationary now. Additionally, the two graphs illustrate that the dispersion of relative unemployment rates is significantly bigger during the second sub-period than during the first one. It seems as if the levels of the series have changed due to a structural break. Finally, note that there is no apparent deterministic time trend in the data.

4 Unit-root tests without structural breaks

Having displayed the series graphically, we turn to a formal characterization of the stochastic behavior. The hypothesis being tested is that relative unemployment rates follow a unit-root process. To set the scene, we first employ a univariate unit-root test without structural breaks. As a next step, we turn to more powerful panel-based unit-root tests. Later on, we extend the analysis to allow for structural breaks.

Figure 2: Relative unemployment rates in West Germany, subperiods 1960-1979 and 1980-2002



.

4.1 Univariate unit-root tests

As explained in Section 2, tests of convergence can be conducted as Dickey-Fuller (1979) type tests (ADF tests) based on the difference between the unemployment rate in federal state i and the unemployment rate for Western Germany:

$$\Delta u_{i,t} = \mu + (\rho - 1)u_{i,t-1} + \sum_{j=1}^{k} \varsigma_j \Delta u_{i,t-j} + \varepsilon_{i,t}, \tag{4}$$

$$u_{i,t} = ur_{i,t} - ur_{Ger,t}.$$

If the series contains a unit root ($\rho = 1$), the proposition for both, absolute and conditional convergence is violated. The alternative hypothesis is that $\rho < 1$, which implies that the series is stationary. Moreover, absolute (or unconditional) convergence implies that the constant term, μ, is insignificant.[6]

The ADF-tests of convergence in relative unemployment rates are reported Table 1, optimal lag lengths, k, have been determined by sequential $t-$tests as suggested by Ng and Perron (1995). It can be seen that there are considerable differences in the time-

[6] We do not include a deterministic time trend in the regressions since a trend is neither compatible with long-term convergence nor apparent in our data.

Table 1: ADF test for relative unemployment rates (without trend)

Augmented Dickey-Fuller (1979) test

Federal State	lags (k)	$\hat{\mu}$	$\hat{\rho}-1$	p-value	Federal State	lags (k)	$\hat{\mu}$	$\hat{\rho}-1$	p-value
BW	1	-.117	-.083	0.402	NRW	5	.099	-.122	0.440
		(.066)	(.047)				(.061)	(.073)	
BY	5	-.087	-.066	0.658	RP	0	-.071	-.402	0.027**
		(.061)	(.054)				(.062)	(.130)	
BRE	0	.283	-.082	0.650	SAAR	3	.321	-.189	0.547
		(.201)	(.065)				(.169)	(.128)	
HH	3	.176	-.081	0.791	SH	0	.278	-.241	0.187
		(.144)	(.091)				(.133)	(.107)	
HE	2	-.044	-.219	0.233	ur_{Ger}	2	.302	-.038	0.729
		(.052)	(.103)				(.191)	(.036)	
NS	0	.105	-.108	0.509					
		(.072)	(.070)						

*,**,*** significant at the 10, 5, and 1 percent levels, respectively.
Standard errors in parentheses.

series properties of relative unemployment rates among the federal states, but the most important result is that for nearly all federal states we cannot reject the null hypothesis of a unit root. The unit root is rejected only for Rheinland-Pfalz.

This means that the ADF tests provide no evidence of stochastic convergence during the period under study. Other studies of convergence often include a deterministic time trend in the ADF regressions. In our setting, the derived results do not depend on the absence or presence of a trend. If we allow for a time trend, results do not change. The series for Rheinland-Pfalz remains (trend) stationary and all other series remain non-stationary.[7]

4.2 Panel unit-root tests

It is well known that unit-root tests such as the ADF test have low power against stationary alternatives in small samples. Panel-based unit-root tests have proven to be more powerful, since they exploit the cross-sectional dimension of the data.

The basic regression for these panel unit-root tests is[8]

$$u_{it} = \rho_i u_{i,t-1} + z'_{i,t}\gamma + \varepsilon_{i,t} \qquad i = 1, ..., N; \; t = 1, ..., T$$

where z_{it} is the deterministic component and ε_{it} is a stationary error term. The set of exogenous regressors z_{it} could be empty, or include a common constant, fixed effects, or fixed effects and a time trend.[9]

The Levin, Lin, and Chu (2002) test (henceforth LLC) assumes that each individual unit in the panel shares the same autoregressive coefficient: $\rho_i = \rho$ for all i. Hence, the power of the single ADF tests is increased not only by pooling the data but also by exploiting a cross-equation parameter restriction on the autoregressive parameters.[10] The null hypothesis of the LLC test states that the relative unemployment series of *each* state contains a unit root, which is tested against the alternative that *all* series are stationary.

The panel regressions of the LLC test include constant terms that reflect fixed effects to control for heterogeneity among cross-sectional units. In our setting, these fixed effects capture stable differences to the national average to which regional unemployment rates

[7] We also tried the Dickey-Fuller GLS test proposed by Elliot, Rothenberg, and Stock (1996). The qualitative results are the same as obtained with conventional ADF tests. Choosing the lag length according to information criteria does neither change the results. Also Phillips and Perron (1988) tests and Kwiatkowski, Phillips, Schmidt, and Shin (1992) (KPSS) tests yield qualitatively similar results. Detailed results are presented in Appendix A.

[8] See Baltagi (2001) for an overview of econometric methods for non-stationary panel data.

[9] In the more general case, when the error disturbances $\varepsilon_{i,t}$ are serially correlated, the serial correlation can be corrected by including lagged terms similar to the ADF procedure.

[10] The LLC test statistic converges more rapidly with respect to the time dimension T than with respect to the cross-section dimension N. Hence, the LLC test is well-suited for our dataset with $N = 10$ and $T = 43$.

Table 2: Levin, Lin, and Chu and Breitung and Meyer tests for a unit root in relative unemployment rates

Levin, Lin, and Chu (2002) test					**Breitung and Meyer (1994) test**					
k	Obs.	$\hat{\rho}-1$	t^{*1}	$P>t^*$	k	Obs.	$\hat{\rho}-1$		t^{*1}	$P>t^*$
							est.[2]	adj.[2]		
0	420	-0.116	-1.850	0.032**	0	420	-0.048	-0.096	-2.510	0.006***
1	410	-0.112	-1.643	0.050**	1	410	-0.040	-0.079	-1.980	0.024**
2	400	-0.117	-1.307	0.096*	2	400	-0.040	-0.081	-1.938	0.026**
3	390	-0.096	-0.013	0.495	3	390	-0.013	-0.026	-0.617	0.269
4	380	-0.101	-0.026	0.490	4	380	-0.017	-0.034	-0.772	0.220

*,**,*** significant at the 10, 5, and 1 percent levels, respectively.
[1] t^* is distributed standard normal under the null.
[2] The estimate of ρ is unbiased under the null but biased under the alternative hypothesis, since $\text{plim}(\hat{\rho}-\rho)=\frac{1-\rho}{2}$. The column 'est.' displays unadjusted estimates which are valid under the null, the column 'adj.' displays the bias-adjusted estimates, which are valid under the alternative hypothesis. The t^*-statistic is computed on the basis of the unadjusted estimates, see Breitung and Meyer (1994).

converge.

Since the LLC test assumes a homogeneous autoregressive coefficient, it has a straightforward economic interpretation, which is its major advantage. We can interpret the autoregressive coefficient as a measure of the average speed of convergence in the sample. The number of years a shock needs to decay by 50% can be computed as $\frac{\ln 0.5}{\ln \rho}$. Knowing the implied half-life is important, because it allows us to compare the results of tests with and without a structural break with respect to the speed of convergence they imply. This interpretational advantage makes the LLC test our preferred testing procedure, but we consider alternative testing procedures to check for robustness.

The left-most columns of Table 2 summarize the results of the LLC test. The inclusion of a time trend does not change the results qualitatively. We can reject the null hypothesis of a unit root safely, if no or only one lag is included to allow for serial correlation in

Table 3: Pooled AR(1) estimation with fixed effects

Fixed-effects regression		
Dependent variable: $(ur_i - ur_{Ger})_t$		
constant	0.073	$(2.74)^{***}$
$(ur_i - ur_{ger})_{t-1}$	0.880	$(36.75)^{***}$
$F(9, 409) = 2.38^{**}$ (indiv. effect is zero)		

*,**,*** significant at the 10, 5, and 1 percent levels, respectively.
The number of observations is 420 (42 years and 10 cross-sectional units).
t-statistics in parentheses.
R^2 within is 0.79.

the error terms. If a second lag is included, we can still reject the null at the 10 percent level. Moreover, the parameter estimate for the autoregressive coefficient does not change substantially across the different specifications. If three or more lags are included, we cannot reject the null hypothesis anymore. Since the univariate ADF tests of the previous section suggest an average optimal lag length of roughly 2, we suppose that the model specification with two lags is most preferable. To corroborate this hypothesis, we consider LLC tests with heterogeneous lag lengths further below and select the number of lags also according to information criteria.

The parameter estimate $(\hat{\rho} - 1) = -0.117$ implies an autoregressive parameter of 0.883. This in turn means that the half-life of a shock to relative unemployment rates is 5.57 years. This seems a moderate degree of persistence.

A testing procedure similar to the LLC test is the one proposed by Breitung and Meyer (1994). This test also assumes a homogeneous autoregressive coefficient.[11] For the Breitung and Meyer (1994) test, a similar pattern emerges as for the LLC test, see the right-hand side of Table 2. If we use less than three lags, we can reject the null hypothesis of non-stationarity. Indeed, the bias-adjusted estimate for ρ is close to the one implied by the LLC test. However, the asymptotic properties of the Breitung and Meyer (1994) test

[11] The Breitung and Meyer (1994) test has been extended to allow for a deterministic time trend by Breitung (2000).

are primarily based on the size of the cross-sectional dimension. Therefore, the test results have to be interpreted with care and the LLC test seems to be preferable.

Having shown that the time series for relative unemployment rates are jointly stationary, we estimate a simple AR(1) fixed-effects model. This allows us to formally test for unconditional convergence by testing the joint significance of the fixed-effects. The fixed-effects estimation is reported in Table 3. The $F-$test that all unit effects are zero is reported in the last row of the table. Since we have to reject the hypothesis that all fixed effects are insignificant, we find no evidence for unconditional convergence of regional unemployment rates.

Both, the LLC test and the Breitung and Meyer test impose the constraint that ρ is homogeneous across cross-sectional units. While this constraint enables us to interpret the test statistics in economic terms, it may be too restrictive from a statistical point of view. Im, Peasaran, and Shin (2003) (henceforth IPS) propose an alternative testing procedure, which allows for heterogeneous ρ_i. This means that the speed of convergence may differ among regions. While the null hypothesis of the IPS test is the same as for the LLC test, the alternative hypothesis is more flexible. It states that at least one of the series is stationary but not necessarily all. The results of the IPS tests are reported in the left-most columns of Table 4. By and large, we find a similar pattern as with the LLC test. Again, the inclusion of a time trend does not alter our findings.

Similar to the IPS test, also the unit-root test by Sarno and Taylor (1998) allows for heterogeneous ρ_i. Additionally, it exploits contemporaneous correlations among the disturbances of the ADF regressions and uses a SUR-estimator for the test. Accordingly, if there is cross-sectional dependence, this estimator gains precision compared to the IPS test. On the basis of the Sarno and Taylor (1998) test, we can reject the null-hypothesis of a unit-root at all lag length considered, see the right-hand side columns of Table 4.

In order to find out whether our results of the panel-based tests are exceedingly sensitive to the choice of a model specification with a homogeneous lag length smaller than 3, we determine the optimal lag lengths for each state separately using three alternative criteria. The first criterion is Ng and Perron's (1995) sequential t-testing method, the second selection method is the Akaike information criterion (AIC) and the third one is the Schwartz criterion (BIC). For brevity, we consider the heterogeneous lag-length specifications only for an LLC test and an IPS test. Results are reported in Table 5. Based on the AIC, we can no longer reject the hypothesis of a unit root. Both, the BIC and sequential $t-$testing allow us to still reject the unit-root hypothesis, though only at a marginal level of significance for the IPS test.

For completeness, we also considered Fisher-type tests as suggested by Maddala and

Table 4: Im, Peasaran, and Shin and Sarno and Taylor tests for a unit root in relative unemployment rates

Im, Pesaran, and Shin (2003) test				**Sarno and Taylor (1998) test**			
Lags	Obs.	$W(\bar{t})$[1]	$P > \bar{t}$	Lags	Obs.	$MADF$	approximate critical value 5%
0	420	-1.958	0.025**	0	420	47.268**	22.744
1	410	-1.506	0.066*	1	410	45.864**	22.974
2	400	-1.615	0.053*	2	400	44.383**	23.218
3	390	-0.003	0.499	3	390	31.705**	23.476
4	380	-0.099	0.461	4	380	36.337**	23.751

*,**,*** significant at the 10, 5, and 1 percent levels, respectively.
[1] $W(\bar{t})$ is distributed standard normal under the null.

Wu (1999) and Choi (2001). Surprisingly, these tests cannot reject the null hypothesis of a unit-root.[12] The result is puzzling insofar as the Fisher-type tests have been designed to alleviate a potential power problem that has been attributed to LLC tests.[13]

Overall and in summary, the panel-based tests show some support for conditional convergence of relative unemployment rates during the period 1960-2002. However, the estimated speed of convergence is slow at best and differences in unemployment rates do not disappear completely over time. If the panel-based tests suggest convergence, then they also suggest that there is a stable distribution of relative regional unemployment rates, i.e. there is conditional convergence only.

Though, the graphical analysis of the time series for relative unemployment rates suggested that there might be a structural break in the means of the series. Hence, our conclusion of sluggish convergence may be premature. If there is a structural break indeed, the estimated degree of persistence will be biased upwards. The interesting question

[12] Detailed results are presented in Appendix A.

[13] However, the gain in power by the Maddala and Wu (1999) test is most pronounced when a time trend is included into the regressions. In fact, Table 1 in Maddala and Wu (1999) suggests that the Maddala and Wu (1999) test may be less powerful than the LLC test for the size of our sample if there is no trend in the data.

Table 5: Levin, Lin, and Chu and Im, Peasaran, and Shin tests for a unit root in relative unemployment rates with heterogeneous lag lengths

Levin, Lin, and Chu (2002) test					**Im, Pesaran, and Shin (2003) test**			
Criterion	Obs.	$\rho - 1$	t^{*1}	$P > t^*$	Criterion	Obs.	$W(\bar{t})^2$	$P > \bar{t}$
Seq. t-tests	401	-0.117	-1.478	0.070*	Seq. t-tests	401	-1.249	0.106
AIC	395	-0.110	-1.226	0.110	AIC	395	-1.173	0.120
BIC	411	-0.109	-1.541	0.062*	BIC	411	-1.339	0.090*

*,**,*** significant at the 10, 5, and 1 percent levels, respectively.
[1] t^* is distributed standard normal under the null.
[2] $W(\bar{t})$ is distributed standard normal under the null.

is whether accounting for the structural break allows us to reject the unit-root hypothesis more clearly and changes the estimated speed of convergence substantially.

5 Unit-root tests with structural breaks

As displayed in Figure 1, the relative unemployment rates for the federal states seem to change permanently about 1980. After 1980, the northern regions, especially the city-states Bremen and Hamburg, exhibit a higher level of unemployment, while the southern states, e.g. Bayern and Baden-Wuerttemberg, experience below average unemployment.

This observation calls for the inclusion of a structural break in the analysis. It also explains why relative unemployment rates are only found to converge conditionally. Absolute convergence implies a zero mean of the relative unemployment series at all times, so that there cannot be structural change. By contrast, conditional convergence implies an equilibrium relationship of regional unemployment rates and the stationarity of their distribution. If the equilibrium relation is non-unique, a major shock may shift the economy from one equilibrium to the other and the relative unemployment rates are only regime-wise stationary. With this regime-wise stationarity, conditional convergence with a structural break implies on the one hand that there is an equilibrium relationship between the unemployment rates of the various states in the absence of major shocks, i.e. regional shocks have no persistent effect. On the other hand, a permanent change of the equilibrium relationship occurs when the regime shifts because of a one-time major shock. To put it simple, if we find evidence for a structural break and convergence, then only very few regional shocks have persistent effects, most of them do not.

Although a theoretical explanation of an apparent level shift is interesting and important (Hansen, 2001), we only try to find the structural break and test for convergence in this paper. A theoretical explanation could for example be based on induced technological change, hysteresis effects, differences in regional specialization, or differences in union density and bargaining power, see Martin (1997) for further examples. We will come back to this issue in Section 6.

5.1 Test procedure

Since we do not specify a structural model for the regime shift, we go back to the univariate time-series approach but extend the model to allow for a one-time level shift. The timing of the level shift, i.e. the structural break, is determined endogenously and data-dependent. This approach follows the testing procedure introduced by Perron (1990), who has shown that conventional ADF tests perform poorly when there is a structural break in the means of the series. Unless the break is accounted for, a conventional unit-root test will falsely suggest non-stationarity of data that is generated by a stationary process which is subject to a structural break. This suggests that the univariate tests presented in Section 4 may have been unable to reject the unit-root hypothesis because of a permanent change in the level of the series about 1980. Similarly, this structural break may also drive the moderate speed of convergence we find on the basis of the panel-based tests.

The original approach proposed by Perron (1990) requires the break date to be known to test for a unit root in the presence of a structural break. Since we do not want to specify a certain break date a priori, we employ the Perron and Vogelsang (1992) test instead which determines the breaking date data-dependent.

Perron and Vogelsang (1992) propose two alternative models to describe the transition of the time series from the old to the new level. The first alternative, labelled 'additive outlier model' (AO), assumes the transition to happen instantaneously after the break has occurred. The second alternative, the 'innovational outlier model' (IO), assumes the break to affect the time series just as temporary shocks to the series. Hence, the adjustment to a new equilibrium occurs slowly over time in this model. The graphical analysis in Section 3 suggests that adjustment after a level shift needs some years to take effect and does not occur instantaneously (see Figure 1). Consequently, the IO model is more appropriate for our data.[14]

[14] This finding is also in line with the general remark of Hansen (2001), who argues that a structural break

The IO model of the Perron and Vogelsang (1992) test can be described as follows. Let T_b denote the date of the break with $1 < T_b < T$, where T is the sample size. The null hypothesis is specified as

$$u_{i,t} = u_{i,t-1} + \psi(L)(e_t + \theta D(TB)_t), \qquad t = 2, ..., T \tag{5}$$

where $\psi(L)$ defines the moving average representation of the ARMA noise function. The dummy variable $D(TB)_t$ is set to 1 if $t = T_b + 1$ and 0 otherwise. The dummy $D(TB)_t$ is a one-off impulse dummy which changes the level of the series after the break by θ under the null hypothesis of a unit root. The long-term impact of the level change is given by $\psi(1)\theta$.

Under the alternative hypothesis of stationarity, the model is represented by

$$u_{i,t} = a + \phi(L)(e_t + \delta DU_t), \qquad t = 2, ..., T \tag{6}$$

where $\phi(L)$ defines the moving average representation of the ARMA noise function under the stationary alternative. The dummy variable DU_t is equal to 1 if $t > T_b$ and 0 otherwise. Hence, the expected value of $u_{i,t}$ becomes $(a + \phi(1)\delta)$ under the stationary alternative in the long run after the break date. As suggested by Perron and Vogelsang (1992), models (5) and (6) can be nested and approximated by the finite-order autoregressive model

$$u_{i,t} = \mu + \delta DU_t + \theta D(TB)_t + \rho u_{i,t-1} + \sum_{j=1}^{k} \varsigma_j \Delta u_{i,t-j} + \varepsilon_{i,t}, \qquad t = k+2, ..., T \tag{7}$$

Similarly to the augmented Dickey-Fuller regression, lags of first-differences, $\Delta u_{i,t-j}$, are included on the right-hand side of the equation. Model (7) can be estimated by OLS. Under the null hypothesis of a unit root, the autoregressive parameter ρ is equal to 1, which implies $\delta = \mu = 0$ if there is no time trend.

Since we do not specify the break date T_b beforehand, we need an empirical strategy to estimate T_b along with the other parameters of (7). For this estimation, there are two options. Under both options, one first performs regression (7) for all possible breaking dates. Then, under the first option, the break date is chosen to minimize the t-statistic on the autoregressive coefficient. In other words, this option selects the break date to provide most evidence against the random walk hypothesis.

The alternative option identifies the break point as the value of T_b that maximizes the

is unlikely to be immediate. Nonetheless, we also tried the AO model, but as expected, its performance turned out to be inferior compared to the IO model. This means that the AO model rejects the null hypothesis in fewer cases.

t-statistic (in absolute terms) on the coefficient associated with the change in the mean, δ. In other words, this option chooses the break date to capture the most significant change in the series.

Perron and Vogelsang (1992) derive asymptotic distributions of the test statistics and finite-sample critical values for typical sample sizes. In order to obtain critical values that correspond exactly to our sample size of $T = 43$ and a maximum lag length of $k_{\max} = 8$, we perform 5000 replications of a Monte-Carlo experiment to simulate the unknown distribution of ρ. There are various procedures to select the appropriate order k of the estimated autoregressions and each procedure influences the distribution of ρ under the null hypothesis. Most prominent procedures are Ng and Perron's (1995) sequential $t-$test, Akaike's information criterion (AIC), and the Schwartz-criterion (BIC).

5.2 Test results

Table 6 summarizes the results of the Perron and Vogelsang (1992) unit-root tests obtained by minimizing the t-statistic on ρ over all possible break points. The augmentation lag length has been determined by using sequential $t-$tests.

In seven out of ten cases, we are able to reject the null hypothesis of a unit root in favor of regime-wise stationarity at least at the 10 percent level of significance. Recall that the univariate unit-root tests without structural breaks rejected the random walk hypothesis only for one federal state. For three of the ten federal states we still cannot reject the null hypothesis of a random walk even after accounting for a structural break. These states are Baden-Wuerttemberg, Niedersachsen, and Schleswig-Holstein. However, the non-rejection seems to be due to a lack of power as point estimates for ρ range from 0.5 to 0.7. The weak power of the test can also be seen if we look at the opposite extreme cases. Although the estimates of ρ for Bremen and Hessen are virtually zero, the test rejects the hypothesis of $\rho = 1$ only at the 5 or even 10 percent level of significance. However, we do not need to worry too much about a potential problem of low power, because we are in fact able to reject the null hypothesis in seven out of ten cases. If the power problem were effective in our sample, a potential way to increase the power would be to exploit the panel dimension again.

The data-dependent choice of the break date mostly coincides with the a priori assumption that the second oil crisis and the following recession had a huge and persistent impact on relative unemployment rates. For all but three series, the chosen break date falls into the period of 1978-1982.

The three states for which the estimated break date falls outside this period are Rheinland-Pfalz, Hessen, and Schleswig-Holstein. For Rheinland-Pfalz, the ADF test without structural break already rejected the unit root. For Schleswig-Holstein, the estimated

Table 6: Perron and Vogelsang unit-root tests, lag length selected by sequential t-tests

Perron and Vogelsang (1992) test

Fed. State	T_b [1]	k [2]	$\hat{\rho}$	$\hat{\delta}$	Fed. State	T_b [1]	k [2]	$\hat{\rho}$	$\hat{\delta}$
BW	80	6	0.49	-0.69	NRW	80	6	0.28**	0.78
			(-4.23)	(-3.76)				(-5.54)	(5.05)
BY	81	5	0.30**	-1.08	RP	70	0	0.38*	-0.36
			(-5.03)	(-4.81)				(-4.72)	(-2.75)
BRE	82	5	0.05**	3.75	SAAR	78	2	-0.41**	2.06
			(-4.99)	(4.78)				(-5.21)	(4.56)
HH	82	1	0.27**	1.75	SH	72	0	0.60	0.39
			(-5.22)	(4.96)				(-3.36)	(2.37)
HE	93	4	0.05*	0.75	ur_{Ger}	79	6	0.65	1.94
			(-4.60)	(3.57)				(-3.67)	(3.53)
NS	78	2	0.69	0.27					
			(-3.51)	(2.30)					

Critical Values[2,3]	1%	2.5%	5%	10%
T_b chosen by min. $t_{\hat{\rho}}$ [1]	-5.61	-5.25	-4.91	-4.53

*,**,*** significant at the 10, 5, and 1 percent levels, respectively. t–statistics in parenthesis.

[1] T_b, k, ρ, θ are obtained by minimizing the t-statistic on ρ.

[2] Lag length k chosen according to a significance test on the last included lag, given a pre-specified maximum of $k = 8$.

[3] Obtained from the empirical distribution of 5000 replications of a Monte Carlo experiment.

break date coincides with the first oil crisis, but the unit root cannot be rejected. Only for Hessen, the break date is hard to interpret. It could be the German reunification of 1989/90 that affects Hessen with a three year time-lag in 1993. But since we cannot give a clear-cut explanation for the break date in economic terms, we may view the test results for Hessen—including the rejection of the unit root—with reservation.

The unemployment rate for Germany as a whole remains non-stationary even after accounting for a structural change in the level. This result is in line with the findings of Papell, Murray, and Ghiblawi (2000).[15]

These results are relatively robust with respect to the two different methods to determine the break point. The two methods do lead to different estimates of break-point and/or a different number of augmentation lags in only two cases. For Niedersachsen, the alternative $\delta-$method estimates the break point to be 1979 instead of 1978 without a change in the qualitative result of non-stationarity. For Rheinland-Pfalz, the $\delta-$method yields 8 augmentation lags and can no longer reject non-stationarity. However, this again reflects low power as we could already reject non-stationarity for Rheinland-Pfalz using the ADF test.

Although Perron and Vogelsang (1992) recommend sequential t-testing, we check the robustness of the results to alternative methods of lag-length selection. Both, the AIC and the BIC tend to choose shorter lag length than sequential t-testing, but the estimated break points remain very similar.[16]

Only for Hessen, we cannot reject the unit root on the basis of an information criterion, but could reject the unit root under sequential t-testing.[17] This confirms our previous warrant concerning the test results for Hessen for which the estimated break date was not intuitive.

Also the results for Bayern and Bremen change under the AIC and BIC, but only quantitatively. The levels of significance pejorate somewhat (from 5% to 10% level). However, this is only due to a marginal change of significance from below 5% to slightly above this level. This can be illustrated by plugging-in the estimated t-values in the simulated distribution obtained by the Monte-Carlo experiment. For the BIC, for example, we get approximate $p-$values of 6.2% and 5.4% for Bayern and Bremen, respectively.

To further test the robustness of our results, we also tried unit-root tests which allow

[15] Papell, Murray, and Ghiblawi (2000) analyze hysteresis in OECD unemployment rates. They adopt unit-root tests with multiple structural breaks and show that the West German unemployment rate is non-stationary.

[16] The tables are available upon request.

[17] The lag-length selection criterion influences the distribution of the t-statistics under the null hypothesis of non-stationarity. Therefore, we have simulated the distributions for each criterion by Monte-Carlo experiments.

Table 7: Half-lifes (in years) of shocks to relative unemployment rates, computed from impulse-response functions based on regression results as reported in Table 6

Federal State	BY	BRE	HH	HE	NRW	RP	SAAR
Half-life	2	1	2	1	3	1	1

Note: Three federal states are omitted, for which the relative unemployment series were found to be non-stationary.

for a break both in the intercept and in the trend (Perron, 1997, Zivot and Andrews, 1992). Allowing for breaks in the time trend provides little additional evidence against the unit-root hypothesis. The unit-root hypothesis cannot be rejected at a higher level of significance because the power of the tests declines when unnecessary breaks are included.

5.3 Speed of convergence

It has been the moderate speed of convergence, which we have inferred from the panel-based unit-root tests, that has motivated us to apply a test which allows for a structural break. To show that the estimated speed of convergence is substantially affected by the structural break, we analyze the half-life of a shock to relative unemployment rates implied by the results of the Perron and Vogelsang (1992) test. This, of course, makes sense only for those regions for which non-stationarity could be rejected. For those states for which the unit-root hypothesis cannot be rejected, shocks have a persistent effect and the implied half-life is infinite.

For those series which are found to be stationary by the Perron and Vogelsang (1992) regressions, we generate a moving-average representation of the estimated autoregressive process that includes the augmentation lags. This moving-average representation is used to compute impulse-response functions and we define the half-life of a shock as the date at which the initial impulse has lost at least half of its effect for the first time.

The estimated half-lifes are reported in Table 7. While the implied half-life is 5.6 years when the LLC test results are used, the half-lifes go down to between 1 and 3 years when we include a structural break. Consequently, measured persistence is substantially biased upwards if the structural break is omitted.

6 Interpretation and discussion

Although it is hard to fix a clear theoretical underpinning for our finding of regime-wise convergence, a potential explanation could be hysteresis.[18] As a theory, hysteresis usually refers to the absolute levels of unemployment and is associated with the existence of multiple equilibria. The multiple equilibria manifest in non-linear, non-stationary behavior of unemployment, which displays a high degree of persistence in turn, e.g. unit-root or close to unit-root behavior.[19] Instead of testing for high persistency, a more direct approach were a test for structural breaks which represent endogenous shifts from one equilibrium to the other, as for example in the 'coconut' model of Diamond (1982).

However, Amable et al. (1991) and Cross (1994) have challenged the latter strategy building on the ideas of Krasnosel'skii and Pokrovskii (1989) and Mayergoyz (1991). They point out that it also depends on the degree of heterogeneity at the micro level, whether hysteretic micro behavior manifests itself in structural breaks at the macro level, or in more general forms of non-linear persistent time-series behavior. If the hysteretic forces are heterogeneous at the micro level, aggregate behavior is smooth, but non-linear and persistent. If there is homogeneity at the micro level, however, hysteresis should result in structural breaks. In any case, hysteresis implies that the distribution of relative unemployment rates is not stable over time.[20]

We do find evidence for a change in equilibrium following the second oil crisis. Our test results overall—low persistency with structural breaks and high persistency without structural breaks—may thus suggest some form of hysteresis driving relative unemployment rates in West Germany, if microeconomic agents are relatively homogeneous with respect to their employment decisions ('strong macroeconomic hysteresis', see Cross, 1994 and Amable et al., 1991).

An alternative explanation for these patterns would be a permanent shift of exogenous parameters that determine the equilibrium (Roed, 1997, p. 394) instead of an endogenous change from one equilibrium to the other as proposed by hysteresis theory. Whether the change in equilibrium forms endogenously or is due to an exogenous and permanent shift of deep parameters can hardly be discriminated on the basis of our univariate analysis.[21]

Although the literature has typically stressed the difference between exogenous change and hysteresis (Roed, 1997, p. 406), both have similar implications for regional policy

[18] We thank an anonymous referee for suggesting this link.

[19] See Blanchard and Summers (1986), Roed (1997), Amable et al. (1994), Amable et al. (2005).

[20] Belke and Göcke (2005) extend this argument to the role of uncertainty in hysteresis. For a survey, see Göcke (2002), which provides an overview of the concepts of hysteresis and their implications for applied economic studies.

[21] A possible way to discriminate would be to analyze the employment behavior at the micro and macro level simultaneously, but this data is not available for the long time-period we want to study.

against the background of our results. Irrespective of how one motivates the permanence of the change in the 1980s structurally, one can expect small government interventions to loose their effect quickly. We find that relative unemployment rates adjust quickly to their equilibrium levels, but in exceptional cases the economy might move from one equilibrium to the other. Consequently, a policy intervention needs to take the form of a substantial intervention or a substantial change in politically set parameters in the case of hysteresis or structural change, respectively.

The question whether hysteresis or structural change is driving our results, hence, determines merely the aim and the means of the substantial policy intervention. It has no influence on the suggested size of the intervention, which has always to be substantial to be effective. We cannot tell which policies are actually likely to reduce relative unemployment dispersion, but most policies that aim at reducing relative unemployment differences are unlikely to make permanent contributions to social welfare because they are simply to small.

One might argue that this conclusion is misleading since we ignore the endogeneity of regional policy. Such policy endogeneity may result in mean-reverting behavior of relative unemployment rates although these rates would be non-stationary in the absence of regional policy. In such setting, regional policy in fact contributes substantially to social welfare by stabilizing the economy and our conclusion above would be just turned on its head.[22]

However, this more optimistic view of regional policy has a hard time to explain why we only find conditional convergence with a structural break. If regional policy were indeed fully effective in reducing the dispersion of unemployment, then one would expect that policy were able to eliminate regional differences completely. In other words, one would have to construct complicated reasoning to justify why the aim of policy should be to obtain an equilibrium dispersion of unemployment in which some states have permanently higher unemployment rates than others. Moreover, this reasoning would need to explain why this dispersion changed in the 1980s. One such explanation could be different costs of regional employment policy. That is, jobs are permanently attracted more cheaply by government intervention in the low unemployment states, e.g. Bayern. However, in this case it must be that the marginal costs of job attraction have changed after the second oil crisis for some states, but not for others. Overall, we find this explanation less intuitive than the simple presumption that regional employment policy cannot permanently attract jobs unless it changes significantly the fundamental economic parameters.

Consequently, this leads us to conjecture that the stabilizing effect of (small but con-

[22] We thank an anonymous referee for pointing this out to us.

stant) policy interventions must be limited even taking into account the problem of policy endogeneity. Nonetheless, we may arrive at a more qualified result with a deeper analysis using more detailed multivariate data or microdata. Yet, for Germany, there is no longitudinal microdata of which the time-span is large enough to cover the second oil crisis. Also at the regional level, there is hardly data that may help to shed more light on the issue of policy endogeneity and covers our sample period from 1960 until 2002.

7 Conclusion

The question of this paper was whether there are forces that lead to convergence in the levels of regional unemployment rates over time. We used German regional data on unemployment aggregated from the Mikrozensus database covering the period 1960-2002 and performed univariate as well as panel unit-root tests to examine the hypothesis of stochastic convergence. On the basis of univariate ADF tests the hypothesis of non-convergence cannot be rejected. But using more powerful panel unit-root tests, we found some evidence for conditional convergence in regional unemployment rates. They converge up to a stable equilibrium distribution. Yet, the panel-based tests imply a moderate speed of convergence at best.

Since the graphical analysis of the series suggested the presence of a shift in the equilibrium differential of regional unemployment rates after the second oil crisis, we extended the convergence tests to allow for such a shift. We employed the univariate unit-root test of Perron and Vogelsang (1992) that includes a level shift in the series analyzed. In contrast to the univariate ADF test, the non-convergence hypothesis could be rejected for seven out of ten federal states. Moreover, the estimated speed of convergence increased substantially in comparison to the results of the panel-based tests. Consequently, regional unemployment rates are found to converge quickly to a constant difference to the national average, but this difference is not the same for the two regimes before and after the second oil crisis.

On the side of the econometric analysis, our paper, like many others, provides once more evidence of the low power of univariate tests in small samples. This problem is especially apparent in the setting with a structural break and we have dealt with it in two ways: including the panel dimension and accounting for the structural break.

The structural break following the second oil crisis reveals the importance of using for our analysis a database that spans a long time-frame. While a shorter series of higher frequency, e.g. monthly data, may be more powerful to quantify the exact speed of convergence in the absence of structural breaks, it would be unable to uncover structural change itself. We have seen the importance of structural breaks for both the empirical results and their interpretation. For example, structural breaks may allow us to discriminate between different types of hysteresis.

In turn, the finding of structural breaks has important implications for policies targeted at regional unemployment rates. If there is regime-wise conditional convergence and fast equilibrium adjustment, then this implies on the one hand that small government interventions loose their effect quickly as unemployment rates adjust back to their equilibrium levels. On the other hand, the result means that large interventions might move the economy from one equilibrium to the other. Hence, policy intervention needs to take the form of a substantial regime shift.

8 References

Amable, B., J. Henry, F. Lordon and R. Topol (1991), 'Strong Hysteresis: An Application to Foreign Trade', *OFCE Working Paper/Document de travail no. 9103*, Observatoire Francais des Conjonctures Economiques, Paris.

Amable, B., J. Henry, F. Lordon and R. Topol (1994), 'Strong Hysteresis versus Zero-Root Dynamics', *Economics Letters* 44, 43-47.

Amable, B., J. Henry, F. Lordon and R. Topol (2005), 'Complex Remanence vs. Simple Persistence: Are Hysteresis and Unit Root Processes Obervationally Equivalent?', *mimeo.*

Armstrong, H. and J. Taylor (2000), *Regional Economics and Policy*, 3rd. edn, Blackwell.

Baltagi, B. H. (2001), *Econometric Analysis of Panel Data*, John Wiley & Sons.

Balz, C. (1999), 'Hysteresis auf europäischen Arbeitsmärkten. Neue empirische Ergebnisse auf der Basis multivariater Einheitswurzeltests', *Konjunkturpolitik* 45, 221-239.

Belke, A. (1996), 'Testing for Unit Roots in West German and U.S. Unemployment Rates: Do 'Great Crashes' Cause Trend Breaks?', *Konjunkturpolitik* 42, 327-360.

Belke, A. and M. Göcke (1996), 'Cointegration and Structural Breaks in German Employment, An Error Correction Interpretation', *Jahrbücher für Nationalökonomie und Statistik* 216/2, 129-152.

Belke, A. and M. Göcke (2005), 'Real Options Effects on Employment: Does Exchange Rate Uncertainty Matter for Aggregation?', *German Economic Review* 6, 185-203.

Bernard, A. and S. Durlauf (1995), 'Convergence in international output', *Journal of Applied Econometrics* 10, 97-108.

Bernard, A. and S. Durlauf (1996), 'Interpreting tests of the convergence hypothesis', *Journal of Econometrics* 71, 161-73.

Blanchard, O. J. and L. Katz (1992), 'Regional Evolutions', *Brookings Papers on Economic Activity* 1, 1-75.

Blanchard, O. J. and L.H. Summers (1989), 'Hysteresis and the European unemployment problem, *NBER Macroeconomics Annual 1986* (ed. S. Fischer), Cambridge, MA: MIT Press.

Breitung, J. and W. Meyer (1994), 'Testing for Unit Roots in Panel Data: Are wages on different bargaining levels cointegrated?', *Applied Economics* 26, 353-361.

Breitung, J. (2000), 'The Local Power of Some Unit Root Tests for Panel Data', in: B. Baltagi (ed.), *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, Advances in Econometrics, Vol. 15, JAI: Amsterdam, 161-178.

Camarero, M. and C. Tamarit (2004), 'Hysteresis vs. natural rate of unemployment: new evidence for OECD countries', *Economics Letters* 84, 413-417.

Carlino, G. and L. Mills (1993), 'Are US regional incomes converging?', *Journal of Monetary Economics* 32, 335-346.

Choi, I. (2001), 'Unit root tests for panel data', Journal of International Money and Finance 20, 249-272.

Cross, R. (1994), 'The Macroeconomic Consequences of Discontinuous Adjustment: Selective Memory of Non-Dominated Extrema', *Scottish Journal of Political Economy* 41, 212-221.

Decressin, J. and A. Fatas (1995), 'Regional Labor Market Dynamics in Europe', *European Economic Review* 39, 1627-1655 .

Diamond, P. A. (1982), 'Aggregate demand management in search equilibrium', *Journal of Political Economy* 90, 881-894.

Dickey, D. and W. Fuller (1979), 'Distribution of the Estimates for Autoregressive Time Series With Unit Root', *Journal of the American Statistical Association* 74, 427-431.

Elliot, G., T. Rothenberg and J. H. Stock (1996), 'Efficient tests for an autoregressive unit root', *Econometrica* 64, 813-836.

Evans, P. and G. Karras (1996), 'Convergence Revisited', *Journal of Monetary Economics* 37, 249-265.

Göcke, M. (2002), 'Various Concepts of Hysteresis Applied in Economics', *Journal of Economic Surveys* 16, 167-188.

Hansen, G. (1991), 'Hysteresis und Arbeitslosigkeit', *Jahrbücher für Nationalökonomie und Statistik* 208/3, 272-298.

Hansen, B. E. (2001), 'The New Econometrics of Structural Change: Dating Breaks in U.S. Labor Productivity', *Journal of Economic Perspectives* 15, 117-128.

Im, K. S., M. H. Pesaran, Y. Shin (2003), 'Testing for Unit Roots in Heterogeneous Panels', *Journal of Econometrics* 115, 53-74.

Krasnosel'skii, M. A. and A. W. Pokrovskii (1989), *Systems with Hysteresis*, Berlin: Springer Verlag.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt and Y. Shin (1992) 'Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?', *Journal of Econometrics* 54, 159-178.

Levin, A., C. F. Lin and C. S. J. Chu (2002), 'Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties', *Journal of Econometrics* 108, 1-24.

Maddala, G. S. and S. Wu (1999), 'A comparative Study of unit Root tests with panel data and a new simple test', *Oxford Bulletin of Economics and Statistics* 61, 631-652.

Marston S. T. (1985), 'Two views of the geographic distribution of unemployment', *Quarterly Journal of Economics* 100, 57-79.

Martin, R. (1997), 'Regional Unemployment Disparities and their Dynamics', *Regional Studies* 31, 237-252.

Mayergoyz, I. D. (1991), *Mathematical Models of Hysteresis*, Berlin: Springer Verlag.

Ng, S. and P. Perron (1995), 'Unit root tests in ARMA models with data dependent methods for the selection of the truncation lag', *Journal of the American Statistical Association* 90, 268-281.

Obstfeld, M. and G. Peri (1998) 'Asymmetric shocks: Regional non-adjustment and fiscal policy', *Economic Policy* 26, 207-259.

Papell, D. H., C. J. Murray and H. Ghiblawi (2000), 'The Structure of Unemployment', *The Review of Economics and Statistics* 82, 309-315.

Perron, P. (1990), 'Testing for a Unit Root in a Time Series Regression with a Changing Mean', *Journal of Business and Economic Statistics* 8, 153-162.

Perron, P. and T. J. Vogelsang (1992), 'Nonstationarity and Level Shifts With an Application to Purchasing Power Parity', *Journal of Business and Economic Statistics* 10, 301-320.

Perron, P. (1997), 'Further evidence on breaking trend functions in macroeconomic variables', *Journal of Econometrics* 80, 355-385.

Phillips, P. C. B. and P. Perron (1988), 'Testing for a unit root in the time series regression', *Biometrika* 75, 335-346.

Quah, D. (1996), 'Empirics for economic growth and convergence', *European Economic Review* 40, 1353-1375.

Reutter, M. (2000), 'Hysteresis in West German Unemployment Reconsidered', *CESifo Working Paper* 240.

Roed, K. (1997), 'Hysteresis in Unemployment', *Journal of Economic Surveys* 11, 389-418.

Rowthorn, R. and A. Glyn (2003), 'Convergence and Stability in US Regional Employment', *mimeo*, Cambridge University.

Sarno and Tylor (1998)

Zivot, E. and D. Andrews (1992), 'Further evidence on the Great Crash, the oil price shock and the unit root hypothesis', *Journal of Business and Economic Statistics* 10, 251-270.